working paper

03 03

Ready, Willing, and Able? Measuring Labour Availability in the UK

by Mark Schweitzer

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 03-03 June 2003

Ready, Willing, and Able? Measuring Labour Availability in the UK

by Mark Schweitzer

The unemployment rate is commonly assumed to measure labour availability, but this ignores the fact that potential workers frequently come from outside the current set of labour market participants, the so-called inactive. The UK Longitudinal Labour Force Survey includes information that can be used to predict impending employment transitions. Using this unique dataset, new measures of labour availability, and indicators based on the more familiar unemployment rate alternatives, can be constructed and are reported here. The micro and macroeconomic performance of these labour force availability measures is compared. Two simplified models, which include several categories of reasons for not working as well as demographic variables, perform particularly well in all of the tests. The implications of these preferred models are further studied in the context of regional regressions and comparisons with alternative data sources. These results together illustrate the important role that some groups of the inactive can play as a source of potential workers.

JEL Classification: J21, J64, E24
Key Words: unemployment incidence, labor force

Mark Schweitzer is at the Federal Reserve Bank of Cleveland and may be reached at mark.e.schweitzer@clev.frb.org or (216) 579-2014. This paper was prepared while the author was working at the Bank of England as part of the Central Bank Institute staff exchange program.
The author thanks Larry Ball, Brian Bell, Ian Bond, Mike Joyce, Stephen Nickell, Jonathan Thomas, and anonymous referees for helpful advice.

1. Introduction

Unemployment, or labour availability more generally, plays a central role in a variety of macroeconomic models, and short-term forecasts of inflation are generally dependent on the level and outlook for the unemployment rate. For this reason central banks frequently cite 'tight labour markets' as a policy concern. Yet the suitability of the unemployment rate as the primary metric for labour availability has not been heavily researched.

The definition of unemployment in most countries follows that of the International Labour Organisation (ILO). As reflected in the United Kingdom's Labour Force Survey (LFS) this includes 'people without a job who were available to start work in the two weeks following their LFS interview and who had looked for work in the four weeks prior to interview or were waiting to start a job they had already obtained.'(1) Unemployment will not be an ideal indicator of labour market availability if the ILO requirements fail to sort individuals well in terms of their willingness to work and/or their likelihood of matching with an employer: for example, if substantial fractions of non-employed who do not meet the criteria to be classified as unemployed are likely to respond similarly in the event of discovering a relevant vacancy.

In practice, many non-working individuals, who do not meet the definition of ILO unemployed, become employed without being recorded as unemployed in the UK Labour Force Survey. These concerns have led to some alternative measures of labour market tightness; for example, the set of alternative definitions reported by the Employment Policy Institute.(2)

---

(1) National Statistics (2001), 'How exactly is unemployment measured?' http://www.nationalstatistics.gov.uk/themes/labour_market/other_features/downloads/unemployment.pdf
(2) Previously published in 'Employment Audit', Employment Policy Institute (1999). The potential usefulness of these indicators for monetary policy is shown in King (1998).

A substantial collection of papers has considered the related, but narrower, question: Are non-employed youths just as likely to take a job as 'unemployed youths'? These papers have broadly been interested in understanding the long-standing issue of high youth unemployment rather than general labour market tightness, but the method could also be relevant in the business cycle context. This question is addressed by tests on whether the status of being 'out of the labour market' or inactive is distinct from unemployment in the sense that transitions between this status and employment are statistically similar.

The studies of US and Canadian data find that the unemployed and inactive are statistically distinct in their transition rates to employment. While the literature concludes that these labour market states should not be amalgamated into the unemployment rate series, in many cases the transition rates from inactivity to employment are predictably higher when certain information is known about the individuals. Notably, the individuals' statement of interest in a job is relevant.

This paper begins with similar statistical tests on UK data. However, our interest in business cycle indicators suggests going an additional step beyond these papers. Even if unemployment and inactivity are statistically distinct, predictable transition rates by grouping or timing may offer valuable macroeconomic information. This paper models the transition process by applying a logit model to the probabilities of transition. The baseline model is estimated with demographic variables and a fully disaggregated set of reasons for inactivity. A key feature of this paper is that it aims to include all of the non-employed within the model, rather than trying to identify a boundary of relevant groups. Similarly, a general set of information on individuals is explored and some demographic variables and non-employment states, not previously identified as important, are shown to be relevant.

While the previous literature has generally been concerned with cross-sectional results, this paper derives a time series of labour availability, in order to evaluate the implications of a broader potential workforce. The limited time period of the analysis, which is constrained by the availability of detailed LFS data, is partially overcome by exploiting the wide regional variation in the United Kingdom. Labour availability differences help to explain regional differences in gross flows to employment and (when specified to account for labour demand) regional wage growth. This is despite the estimation being limited to a time when both the wage growth and participation patterns have been difficult to explain.(3)

Before turning to the analytical parts of the paper, we discuss the relevant labour market literature from a business cycle perspective. A descriptive model of the labour market with different labour market states is provided to clarify the labour market features needed for these estimates to be relevant. We also review the UK Labour Force Survey data and the variety of information it provides for this issue. The nature of the data allows us to make certain comparisons that were not available to the previous papers.

2. Literature review

The literature on transition rates from other labour market states into employment was initiated by Clark and Summers (1979), who suggested that transition rates for American teenagers from inactivity to employment is no different from the rate from unemployment to employment. This conclusion, and its implications on search-theoretic models, led to statistical analyses of the equivalence of unemployment to employment and inactivity to employment transition rates. Three papers stand out in this literature, although the last is perhaps the most relevant to this paper.

(3) Nickell (2001) stressed these anomalies in a recent speech to the Society of Business Economists.

Flinn and Heckman (1983) put the observation made by Clark and Summers to the test, with an eye toward supporting the search-theoretic conception of unemployment for youths. Their tests are the basis of the recent literature on this topic. Flinn and Heckman use a duration of status approach to conclude that transition rates from inactivity to employment and unemployment to employment are statistically distinct even if similar, in a National Longitudinal Survey of Youth sample. Tano (1991) extends this analysis in that he uses Current Population Survey-based Gross Flow data for the analysis, which makes the analysis directly comparable to the unemployment rate but loses simplicity due to the lack of directly observed, individual transitions in his data source.

Jones and Riddell (1999) based their work on Canadian labour market status data that reveals the respondents' desire for work as well as alternative search methods. This expands the set of possible transitions that can be compared in the direction of the data available in the UK. Desire for employment is shown to be useful in predicting subsequent employment. This supports alternative specifications along the lines explored in this paper, although the exact specification the authors would recommend is left unclear. Unfortunately, the analysis in previous papers in this literature generally ends when the alternative statuses are shown to be different from the unemployed. Jones and Riddell move part way from the narrow definition, but limit their analysis to clear alternatives, which are relevant in US/Canada comparisons.

In the United Kingdom, Gregg and Wadsworth (1998) considered the same question with a few categories of the inactive as identified in the LFS. Their work showed that the category of those searching but not currently available for work was largely equivalent to the medium-term unemployed (durations from 6 - 12 months). This paper inspired non-employment index developed by the Bank of England, which was introduced in the August 1999 *Inflation Report*.

This paper moves beyond this work by offering a more general approach that yields easy comparisons between alternative summaries and by focusing on the macroeconomic implications of labour availability.

Shiskin (1976), a commissioner of the US Bureau of Labor Statistics (BLS), was an early proponent of reporting a few alternative unemployment indices. These indices have been published for many countries including the United States as part of their official statistics. Some of these measures focused on labour availability, but other measures were also concerned with measuring the social consequences of unemployment; so it is far from clear that these measures should perform well when focused only on availability. The Employment Policy Institute publishes alternative UK unemployment indices that are similar to the BLS indices, but have been modified to be more consistent with UK data.

Overall, the literature has either offered a framework for testing the behavioural equivalence of alternative labour market statuses without a clear implied measure or alternative measures without a clear interpretation for labour availability. Later papers get closer to looking for alternative measures of labour availability by pointing out the information that might be relevant, but they offer no means of contrasting alternative summaries and do not explore what these measures might contribute in a time series setting.

3. Labour availability with several non-employed states

To illustrate the decisions that underlie alternative non-employment states and the representation in statistical form, consider a simple model with non-employed people, a statistical agency, and a firm. The individual compares a notional wage w, which depends on personal characteristics x and aggregate phenomenon y with the notional value of alternatives u, which are also time varying and idiosyncratic. Formal job search is one of the options and yields a higher probability of finding a job and lower value of non-working time (set to zero for simplicity). By construction, all other options yield a lower probability of employment, but higher values of non-working time.

Which alternative statuses of non-employment can exist in equilibrium? Statuses (o_t^j) that are optimal for at least one individual exist, although optimality would depend on wage and probability parameters in part determined by employers. The problem could be formalised as follows:

$$V_{t+1}(o_t, x_{it}, y_t) = \max \begin{Bmatrix} p_1(o^1{}_t, x_{it}, y_t) w_{it}(x_{it}, y_t) + (1 - p_1(o^1{}_t, x_{it}, y_t)) u_{1i}(x_{it}, y_t) + V_{t+1}(o^1{}_t, x_{it}, y_t) \\ \vdots \\ p_j(o^j{}_t, x_{it}, y_t) w_{it}(x_{it}, y_t) + (1 - p_j(o^j{}_t, x_{it}, y_t)) u_{1i}(x_{it}, y_t) + V_{t+1}(o^j{}_t, x_{it}, y_t) \end{Bmatrix}$$

This value equation expresses the utility *V( )* of choosing an option (o_t^j). In this case it is left quite general with (time-varying) individual characteristics x_{it} and aggregate conditions y_t influencing all components of the model. The probability p_j is also a function of individuals' characteristics and aggregate situation which are only constrained to lie within the interval [0,1), so some states may be consistent with never taking a job. More detailed assumptions could be used to tighten this model to allow more definite conclusions on the existence of unique options, but for the purposes of this paper we will rely on revealed preferences arguments and the empirical evidence shown later. The model merely shows that a rigorous treatment of the decision problem would not preclude the existence of several statuses of non-employment.

Would firms hire someone who was not actively searching for a job? Firms can be thought of as minimising their costs of filling jobs. If appropriate individuals would simply line up outside a firm to see if they had an appropriate job, then that would be the only form of search sustainable for a cost-minimising firm. Instead, firms apply a number of recruiting technologies to attract workers, ranging from listing jobs with public agencies to posting signs on the door. One interesting variant seen in some markets are fast-food restaurants printing job applications on placemats to make sure that customers are aware of their openings. This variety of recruiting strategies seen in labour market suggest a cost minimisation problem where the probability of getting a successful application is weighted against the cost of soliciting applications by a particular mechanism.

For the p_j to be positive for non-employment statuses that do not involve formal search, firms must sometimes apply recruiting strategies that do not require the individual to search for employment. Employers directly asking individuals if they are interested in a job is the type of low effort search we have in mind; for example, the pub owner asking a man at the pub if he would like a job behind the bar. Again a more formal modelling approach could identify what conditions are needed for existence of recruiting strategies, but we will rely again on revealed preferences suggesting the possibility of multiple channels for matching.

All of this reliance on revealed preferences places the 'proof' firmly in the hands of the statistical agency. The statistical agency asks individuals who are not working whether they are looking for and available for work; and, if not, what they are doing. Individuals are assured that their answers are kept confidential to encourage them to answer truthfully. In the UK case, the Office for National Statistics collects statuses on the non-employed in the Labour Force Survey quarterly. This information could potentially identify the probabilities associated with transitioning from a non-employment state to employment.

There are a number of potential problems that could interfere with correct identification:

the agency does not know in advance which information on states indicates distinct statuses; the agency may also be given incorrect information; and the agency can only ask infrequently. The first two problems potentially group distinct types of the non-employed together, averaging together their p_j. In the extreme, this would result in a single probability of transition that would vary little between groups identified by the survey. The final problem is more complicated, because we might miss transitions between states. What this does to the measured probabilities is complicated, because it depends on what the transition rates are between non-employed states. But in the extreme, if every transition between non-employed states was rapid and random then the estimated transition rates would converge to the population average. One possible outcome that cannot be ruled out is that all individuals who get a job enter a momentary period of formal search at least momentarily before becoming employed, but this requires potentially substantial transition rates from inactive statuses to formal search.[(4)]

If the data are meaningful and the model is appropriate, a range of transition rates should be identified, from formal search at the top to very low transitions rates for activities where the individual is largely unable or unwilling to do most any type of work. This result follows from individuals' making decisions about their activities that meaningfully alter their probability of finding a job, employers using recruiting mechanisms that reach people who are not searching and finally the statistical agency asking questions about status that are informative. Without these three conditions applying, transition rates should be approximately random for any grouping of the non-employed.

If the model applies, these probabilities may also vary in predictable and informative ways to business cycle conditions. But that would depend on how aggregate influences enter

[(4)] This is because any probability of direct transition from an inactive status to employment could be matched by into the individuals' decisions. In particular, the number transitioning from inactive statuses has to vary distinctly from the numbers flowing from unemployment. Otherwise, these transitions would be picked up in any time series regression by the unemployment rate. In part these questions are empirical but in practice the business cycle variables one would expect to be associated with changes in labour availability depend on the macroeconomic model favoured. The options here are too plentiful to be reviewed, but the decision problems sketched here should be conceptually consistent with most models that rely on optimising agents.

4. Data

The Labour Force Survey (LFS) data are exceptionally detailed in the information that is provided about the methods of employment search and current activities of the non-employed, which could be used to generate sub-populations that are potentially available for work. In addition, explanatory variables in the model would include basic demographic variables, because these variables are also likely to influence peoples' willingness to work. The presence of children in the home is also potentially informative on participation, as demonstrated by the female labour supply literature.

Critical to this paper is the longitudinal version of the LFS, which exploits the fact that surveyed households continue in the survey for five consecutive quarters. The Office for National Statistics (ONS) matches the data for consecutive quarters using unique individual identifiers.[5] The ONS also produce population weights for this dataset to estimate UK-wide summaries, which adjust for non-response and attrition at the household or individuals level. Attrition is a substantial problem in the longitudinal data, because household or individuals that

---

some unobserved transition rate from inactive to searching times the probability of finding a job when searching.

[5] This makes this dataset a more accurate panel than the matched Current Population Survey datasets used in the United States, which are matched by researchers based on reported characteristics after the dataset has met Census

move from their residences are not followed. This dataset is more fully described in National Statistics (2000).

The primary focus of this analysis is on transitions to employment made by the inactive and unemployed. Given data from a representative sample of potential workers in two successive quarters these transition rates are observable. Table 1 shows these quarterly transition rates from the LFS survey for the full set of reasons for inactivity, covering the period from Spring 1993 to Winter 1999. In addition, unemployed workers are separated into the conventional split between short and long-term unemployed.

Table 1 reveals that there is indeed information in the reasons that people cite for their inactivity that could potentially be used to predict the probability of their entering into employment. As indicated by the transition rate shown in the first column of Table 1, some groups of the inactive are much more likely to become employed in the following quarter than the typical unemployed person. Notably, several categories of individuals who are seeking work, but are not currently available and people who are waiting for the results of an application. These raw transition rates provide preliminary evidence that these individuals should be considered similar to an unemployed person, from the perspective of potential labour supply. However, these are narrow sections of the population (as indicated by the third column of Table 1), summing to only 1.5% of the non-working population of working age. Despite this, they account for 6.4% of those who successfully transitioned from non-employment to employment on average over 1993 to 1999.

Other categories have lower transition rates, but still contribute significantly to gross flows into employment. Indeed, the transition rates for students of all varieties are quite high:

Bureau confidentiality requirements.

ranging from 13.2% to 27.1% for individuals citing being a student as their reason for inactivity. Students were 10.6% of the non-employed population in the sample period. These figures combine to yield 23.4% of transitions to employment, which substantially exceeds transitions from the long-term unemployed. There is every reason to expect that student transitions may be less affected by the business cycle or labour market tightness than the unemployed, but—even so—it is not at all clear that this substantial fraction of the non-employed can be ignored in a business cycle treatment of the labour market.

5. How similar are transition rates?

The underlying statistical model used here follows Flinn and Heckman (1983) in focusing on transition rates to evaluate the reliability of distinctions between unemployment and inactivity. The UK data allow the simpler direct transition probability specification used by Jones and Riddell (1999), which is useful in this case because the duration of inactive status is unknown. The focus of this paper is the availability for work based on the individual's labour market status (U or I) and the reasons for their inactivity. To simplify the analysis, we ignore transitions out of employment and transfers between non-working states.

The UK data allow individuals to explain their labour market inactivity according to their primary alternative activity (eg student, looking after family, sick), whether they were available to work, and finally whether they desired work. This results in a combined categorisation of the inactive into 24 mutually exclusive categories (which are identified below by I6-I29 to accord with the coding in the data). This paper does not *a priori* ignore any of these distinctions, so the simplified transition matrix becomes

$$\mathbf{P} = \begin{pmatrix} p_{E,E} & p_{E,N} \\ p_{U,E} & p_{U,N} \\ p_{I6,E} & p_{I6,N} \\ \vdots & \vdots \\ p_{I29,E} & p_{I29,N} \end{pmatrix}$$

where $p_{A,B}$ represents the probability of transition from state A to B. E represents employment, U conventionally defined unemployment and N all forms of non-employment. So $p_{U,E}$ is the probability of moving from unemployment to employment, while $p_{I6,N}$, represents the probability of moving from inactive category 6 (students, unavailable to start, but seeking work) to all forms of non-employment.(6) The specifics of these categories can be seen in Table 1.

This results in a single set of tests, because the columns are linearly dependent by construction. The second column could be expanded to address each of the labour market states considered in the paper, but this would not substantially alter the equivalence results. Any rejections of a row being distinct in terms of employment outcome is sufficient for general rejection of equivalence of rows for the full matrix, but it is not a necessary condition because each column implies its own tests.

The estimation of these probabilities is both simplified and made more flexible by assuming a logistic distribution for individual *i's* decisions given a state variable z_i. This simplifies the estimation of the full range of transition probabilities by allowing use of standard logit models for the calculation of matrix **P**. Summarising the non-overlapping initial states with a vector of dummy variables s (excluding one category and including an intercept α) results in a coefficient vector γ such that for any person i in category k

$$F(z_i = k) \equiv F(\mathbf{s_i}\boldsymbol{\gamma} + \alpha) = F(\gamma_k + \alpha) = p_{k,E} \text{ , for each labour force status } k.$$

(6) Inactivity categories are numbered according to the ONS labour market status variable INECACA.

This approach is more flexible in that it allows a simple mechanism for other factors to affect the transition probabilities. Regardless of their reason for inactivity, individuals may be more or less likely to begin working based on their age, sex, or education. At the same time, labour force statuses may be correlated with these attributes, which suggests controlling for these characteristics so that their influences can be estimated separately. These conditional estimates are easy to calculate, if one assumes that these factors operate linearly on the underlying index of individual interest in working.

$$F(z_i = k | \mathbf{x_i} = \tilde{\mathbf{x}}) \equiv F(\mathbf{x_i}\boldsymbol{\beta} + \mathbf{s_i}\boldsymbol{\gamma}' + \alpha' | \mathbf{x_i} = \tilde{\mathbf{x}}) = F(\tilde{\mathbf{x}}_\mathbf{i}\boldsymbol{\beta} + \gamma_k' + \alpha') = \Pr\left(E | k, \mathbf{x_i} = \tilde{\mathbf{x}}\right)$$

The primes denote the fact that including the additional covariates will change the parameter estimates on the status variables, while $\tilde{\mathbf{x}}$ refers to any vector of conditioning characteristics. A variety of transition probability tests can be implemented via simple coefficient comparison tests. More generally, likelihood ratio tests can be formed for any alternative that can be summarised as a restriction on the coefficient vector.

This framework is easily extended to allow for separate unemployment statuses based on the time spent without a job. While the literature on transitions has previously addressed this issue via duration models, this paper applies a simpler alternative that addresses the duration distinctions as discrete shifts in the probability of transition. This approach can be implemented by simply augmenting the **P** matrix for more states by disaggregating the duration-dependent category—in this case, unemployment.[7] Given the focus of this paper on non-employment states where duration is not well measured, this approach seems preferable to applying duration analysis to the problem, because it avoids additional functional form assumptions.

[7] This technique imposes a constant hazard rate within the ranges of duration included in the model. While this restriction might appear substantial, Gilleskie and Mroz (2000) show that this approach is capable of very good nonparametric approximations to the underlying density function that is the focus of the hazard rate analysis.

The raw transition rates support further analysis, but statistical tests are needed to establish which of these 'reason' variables are independent of other explanatory variables and how distinct they really are. They could, for example, be dominated by other equally observable factors like age, education, or sex. Equally, transition rates could be statistically unreliable, or perfectly correlated across time with a more commonly revealed summary statistic like the unemployment rate. Finally, even if all these variables were shown to be important, a briefer summary of these factors would be useful in a policy context.

Table 3 compares the marginal probability information implied by alternative logit models of the probability of moving into employment. The results are reported as average marginal probabilities, which show the expected change in probability from a marginal change in a given explanatory variable based on the values of the other explanatory variables.(8)

Note that the estimates are normalised to the excluded category, which in each case was chosen to be a large and low-transition category such as early retirees. The baseline transition rate is also estimated to replicate transition rates in the third to fourth quarters, again based on including seasonal dummy variables (quarters 1, 2, and 4). Variations in the comparison groups means that the results are not directly comparable in levels of the marginal probabilities,

---

(8) This requires information on the values of the other explanatory variables. The more common approach is to report the marginal effects at the means of the explanatory variables. The average marginal approach instead reports the average of the marginal probabilities for each observation, which avoids problems due to the inclusion of dummy variables where the mean of the data cannot occur. For the *k*-th continuous variable the formula is

$$\overline{M}_k = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial F(\mathbf{x}_i\hat{\boldsymbol{\beta}})}{\partial x^k} = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}^k f(\mathbf{x}_i\hat{\boldsymbol{\beta}})$$

where $\mathbf{x}_i\hat{\boldsymbol{\beta}}$ is the index value for person i drawn from the full sample of N records. For the *k*-th binary variable the reported statistic is

$$\overline{M}^k = \frac{1}{N}\sum_{i=1}^{N}\frac{\Delta F(\mathbf{x}_i\hat{\boldsymbol{\beta}})}{\Delta x^k} = \frac{1}{N}\sum_{i=1}^{N}\left(F(\mathbf{x}_i^{k=1}\hat{\boldsymbol{\beta}}) - F(\mathbf{x}_i^{k=0}\hat{\boldsymbol{\beta}})\right)$$

where $\mathbf{x}_i^{k=1}\hat{\boldsymbol{\beta}}$ is the index value for person i when dummy variable k is set to one, while $\mathbf{x}_i^{k=0}\hat{\boldsymbol{\beta}}$ is the same with the dummy variable set to zero.

although the overall predictions are similar once the intercept is accounted for.

For ease of comparison, column 1 of Table 3 repeats the raw transition rates shown in Table 1. The second column adds controls for the seasonal pattern in the data. This is important, because the seasonal pattern is very marked and might distort the transition rates of some seasonal groups. Because all subsequent models include seasonal controls, this is the relevant benchmark for whether a model adds to our understanding. Not all of the standard errors for these coefficients are shown, to keep the table manageable, but the standard errors on the reasons variables vary from less than 0.002 to 0.038 making most of the coefficients significantly greater than zero at all normal significance levels.

The hypotheses that we are actually interested in are comparisons to the unemployment rate. For example, we would like to know whether the transition rate from the inactive category '*students seeking employment but not available*' is comparable to being unemployed, given some control variables. The test for equivalence to the short-term or long-term unemployed is simply a test for coefficient equality, but the test for the overall status of being unemployed is most easily completed via an auxiliary regression where being unemployed replaces the separate short-term and long-term unemployment terms. It is also relevant from a practical standpoint when the estimated transition probability is higher for a given group of the inactive than for the unemployed. The tests included in Table 3 in order of transition rate (highest to lowest) are:

- ♠♠ Transition rate given other control variables significantly greater than the short-term unemployed (95% confidence).
- ♠ Transition rate given other control variables insignificantly different from the short-term unemployed (95% confidence).
- ♥♥ Transition rate given other control variables significantly greater than the unemployed (95% confidence).
- ♥ Transition rate given other control variables insignificantly different from the unemployed (95% confidence).
- ♦♦ Transition rate given other control variables significantly greater than the long-term

unemployed (95% confidence).

♦ Transition rate given other control variables insignificantly different from the long-term unemployed (95% confidence).

Only the test for the highest transition rate is shown in Table 3. For example, two diamonds could be applied to any of the higher categories given the empirical transition rates of the unemployed and the accuracy of these estimates, but if a given inactive group qualifies for a heart then no diamonds are shown. The presence of *any* of these symbols calls into question treating the unemployment rate as the fraction of the population that is available for work, although the tests in the existing literature focus on the hypothesis indicated by hearts.

Unfortunately, information on the duration of inactive states is not available, but evidence drawn from the unemployed shows that the likelihood of becoming inactive rises with the duration of unemployment. In addition, many reasons cited for inactivity are long-term issues. Both of these tendencies would appear to increase the typical interval from last employment of the inactive relative to the unemployed, making long-term unemployment the most reasonable comparison group.

Applying these criteria to the simplest model (estimates with only seasonal controls) indicates that 13 of 26 categories of the inactive are in some sense comparable to the unemployed in their availability. If the higher standard of matching the transition rates of the *overall* unemployed population is applied, 10 categories of the inactive are just as relevant from a labour supply perspective.

Column 3 adds control variables available at the individual level. The standard human capital variables of age, education, and sex are always available for each LFS interviewee. The inclusion of these controls is supported by the coefficient estimates on these variables, which are always significant with the possible exception of the individual's sex. In any case, it appears that there is important, independent information both in individuals' characteristics and in their stated reasons for non-employment.

Two broader categories of controls were also tested: more complete gender controls and family status variables. Both of these alternative sets of controls are suggested by concerns that women's responses to the labour market are distinct, even after controlling for citing family responsibilities as a reason for non-employment. The first option is to allow for a differential response for women, potentially accounting for different behaviour in child-bearing years and any difference in the age-earnings profile of men and women. When a three-term polynomial expansion on women's age is added to the regression in place of the female dummy variable shown in the third column of Table 3, only two of these terms are individually significant and the coefficients are unchanged to the second decimal point. For a sample size of 438,088, these changes are trivial compared with the other variables in the model, so the baseline estimates exclude these terms.

Does adding demographic controls alter the transition rate estimates? Only two reasons for inactivity, which were either statistically indistinguishable or had higher transition rates than the long-term unemployed, fall off the list when additional covariates are included. Notably, it is the student categories that are most affected by the significant age and schooling patterns.

In any of the specifications, virtually all categories of individuals classified as inactive but seeking work (on the basis of their availability) have relatively high transition rates. The exception to this rule is people who cite long-term sickness as their reason for not working. These individuals may qualify for benefits based on a disability, enabling them to stay out of the active population. The other categories of the inactive with high transition to employment are (inexplicably) those who cite little reason for their inactivity [9] and (sensibly) people waiting for the outcome of a job application.

These results can be compared with those of Jones and Riddell (1999) for Canadian data. They find that while wanting a job is a positive factor in transitions, it is not equivalent to being 'unemployed'. In part, the difference in findings is due to the more detailed delineations of the inactive available in UK data. This allows us to conclude that several categories of the inactive should in fact be treated as equivalent to the unemployed.

Adding these inactive categories to the unemployed would however be an overly simplistic reaction to the results presented here. They point strongly to varying, but measurable, degrees of availability among all groups of the non-employed. There is no obvious reason why the analysis of the potentially available workforce should stop at the transition rates of the long-term unemployed, in the sense that this too is an arbitrary benchmark.

6. What makes a good availability summary

While these statistical tests conclude that a number of categories of the non-employed are not equivalent to the unemployed, the transition rates may still indicate that they are relevant to overall labour availability, particularly if they represent a large population. Nonetheless, the number and variety of labour force status categories available in the UK data does not make it easy to produce a complete measure. The fact that some of the categories are small and only marginally different suggests the possibility that the substance of the data can be summarised without much loss of information.

6.1 Summarising availability

Relative probabilities can be used to construct a hedonic index of availability. The

[9] These individuals offered no reason for their economic inactivity, but did answer that they did not want a job.

relative probabilities estimated for a given statistical model are used to weight the categories of the non-employed:

$$LA_t = \sum_{i \in \{\text{non - employed at t}\}} wgt_i F(\mathbf{x_i}\boldsymbol{\beta} + \mathbf{s_i}\boldsymbol{\gamma}' + \alpha)$$

where the wgt_i is the LLFS weight corrected for sample losses in the longitudinal form.

This formulation is equal to the expected number of transitions based on average transitions rates over the full sample period and the characteristics of the workforce in a given quarter. If the number of non-employed or the relative numbers of high transition rate groups decline, then the expected number of transitions will decline.

Several options for summarising labour availability have been offered in the literature. Table 2 shows how some of these alternative summaries rank labour force availability information present in the 23 LFS categories. There are measures based on the alternative unemployment indicators based on Shiskin's work: namely, the series published for the US by the BLS; and the unemployment rate modifications previously published by the Employment Policy Institute. In addition, there is the split featured in Jones and Riddell's work; and the Bank of England non-employment index, which is in part based on the work of Gregg and Wadsworth.

All of these summaries are shown in Table 4, along with a measure based on the ordering and grouping of the outcomes from Table 1. This definition involves making arbitrary rules on splitting the non-employed into groups based on the raw probability of transitioning to employment (as shown in column 1), with an eye towards grouping the data around a few large categories that are intuitive and well understood, like 'taking care of family' or 'students'. The information is clearly significant and results in clean splits between the ordinal reason variable, despite the addition of control variables. Relative to the other summaries this one should have some advantage in being directly based on the UK data and transition rates from the LFS categories.

The final column of Table 3 shows the effect of summarising the reasons based on the 'ordinal reason' definition offered in Table 2. The standard errors for these estimates are shown in parentheses. While these standard errors are not suitable for testing the hypothesis that any of the detailed reasons shown in the previous columns could be subsumed, these standard errors indicate that sampling uncertainty is too low for a likelihood ratio analysis to consistently recommend combining categories. The likelihood ratio for this particular summary (versus the baseline shown in the column 3) is 2,148, which yields a χ^2 value well in excess of all standard significance levels.[10]

The other summary estimates are reported along with the ordinal reasons regression in Table 4 to allow comparisons of the alternatives. In each of these cases, the summary measures are designed to indicate information on the non-employed according to groupings of categories. When a measure reports alternative indicators, these are converted into dummy variables for the additional components. The estimates reported here keep the individual characteristics as control variables and ignore (in favour of direct estimation) any weighting scheme implicit in the measure. The results shown are therefore not a product of the published summaries, but instead are designed to give each summary the best possible chance to successfully describe the pattern of transitions in the data.

The overall patterns continue to hold, notably that—along with the individual's characteristics—information on the reasons that an individual cites for their non-employment is highly informative. Remembering that all of the marginal rates are relative to the excluded

[10] Differences in likelihood ratios over 20 are statistically significant at the 99% level. Most any test of a nested

category, which varies in these specifications, the coefficients on the statuses are similar in their sizes. Each of these summaries involves a larger reference (excluded) category compared to the baseline model, which had early retirees as its reference group. Expanding the reference category of workers from the inactive shifts the relative probabilities for all reasons but generally leaves estimates on the individual characteristics largely unchanged.(11)

6.2 Goodness of fit

How well do these alternative summaries fit the data? This issue is explored via several alternative goodness-of-fit measures in Table 5. The first column of Table 5 shows the results of the simplest Likelihood Ratio (LR) test: for the hypothesis that the predicted probability of entering employment is unaltered by summarising (reducing the variables used in) the detailed specification for the reasons of inactivity (estimates shown in the third column of Table 4). LR tests were also calculated, but not included in the table, for the closer pairings of the models when nested. These tests also uniformly rejected the simpler model. Overall, LR tests continue to favour the inclusion of detailed reason variables.

The uniformity of the negative results from LR tests suggests using alternative measures of the fit, which are shown in the other columns of Table 5. An obvious possibility in this context is the portion of employment predictions that are classified correctly. This and other measures of fit for logistic models are defined in the appendix. The problem with this indicator is that it can fail to evaluate the model predictions when the events are rare. This is because, when transition rates are low no model prediction needs to surpass the 0.5 threshold needed to predict a

---

model in this analysis vastly exceeds these thresholds.

(11) The exceptions were the small, uncertain female coefficient and the variable for having a college degree, which shows a marked decline in the ordinal reasons specification. It is interesting that once reasons for inactivity are accounted for the male/female distinction does not matter very much, despite the fact that most published summaries of labour market conditions make this split. The education coefficients may shift due to the ordinal reasons summary including a category largely based on students, which would suggest that the impact of education might be

transition, even though the model may well be able to sort between groups in their relative probabilities. While transitions to employment are not rare they are infrequent enough that most predicted probabilities of the model are well below 0.5.[12]

This problem is evident in the results of this measure shown in Table 5. In several cases, alternative model specifications correctly predict exactly 90.4% of the transitions. That seems high, but this occurs because in these cases the index never exceeds 0.5 meaning that all non-transition cases are correctly predicted, but no other cases. It is difficult to exceed this figure, because it requires identification of groups of individuals whose transition probabilities exceed 50%. Only the base model and the ordinal reasons variables actually improve forecasts on this basis, with the base model doing better.

This problem with classification-based evaluations has led to alternative prediction-based goodness-of-fit measures. Ben-Akiva and Lerman (1985) suggest a measure based on the average probability of correct prediction. This measure (R_{BL}^2) accounts for the frequencies of true and false predictions made for groups on a metric similar to the traditional R^2: perfect prediction result in a 1. However, this measure has been criticised for not discriminating well in models with unbalanced outcomes, because many correct predictions of the primary outcome (while easy to get from a model) are assigned most of the weight in this measure. Cramer (1999) recommends an alternative that more heavily penalises incorrect predictions of the primary outcome when outcomes are unbalanced. The typical forecast error in each outcome is given equal weight regardless of the relative frequencies of the outcomes.

The value of either these alternative measures is evident in Table 5. Both of these

primarily on those who are or recently were students.

(12) Reducing the threshold inevitably results in over prediction of the rare outcome, although comparisons between models may be more accurate at these levels.

statistics are more informative in their ordering of the models, but (in this case) the measures tend to agree on the ordering of the fit of the models. In both cases, the ordinal reasons model and Bank of England non-employment index lead the set of reduced specifications. Cramer's λ tends to show larger losses than R^2_{BL} for weaker-performing specifications, which reflects the higher penalty for false positives given the relative rarity of transitions. These rules suggest that we can do better than focusing on the unemployment rate as the only measure of labour availability. Likewise, these measures also strongly support using individual characteristics to improve model projections—including them improves predictions by about two thirds.

Finally, we consider the root mean squared error (RMSE) of the quarterly estimate of the flow. This measure allows for errors at an individual level to cancel out as long as these errors are not systematic in a quarter. This measure gets most directly at the issue of forecasting flows for a time series, but may miss early signs of model misspecification that would be averaged out of the time series but be evident in the individual data.

This measure deviates the most in its ranking (shown in Table 5) from the others, because by design it ignores individual outcomes. The base model and all of its variants perform poorly. Simpler specifications other than just the unemployment rate do far better. This is probably the analogue of over-fitting in a panel data set. Statistically significant factors that apply to narrow groups appear to hurt the aggregate performance of the base model. The favoured summaries are the Bank of England non-employment index, the Jones and Riddell summary and the ordinal reason model, in that order. The US Bureau of Labor Statistics model also does well.

One of the surprising conclusions from the RMSE analysis is that individual characteristics continue to aid in prediction. This is surprising because many demographic factors change relatively slowly and might be not expected to augment the predictions in aggregate while being a strong influence at the individual level. In all of the comparisons of models with and without individual characteristics the inclusive models fit the aggregate trends substantially better.

Overall, the ordinal reason variables proposed in this paper or the Bank of England non-employment summary variables with individual level demographics do well on all of these fit evaluations. The Jones and Riddell variables, along with the US Bureau of Labor Statistics, do not perform well at the individual level, while the base measure does not perform well at the aggregate level. The intersection of those that are among the best fit by all criteria has no formal justification, but seems reasonable as an indicator of model 'robustness'.

6.3 Model performance over the sample period

While the ordinal reason variables perform well over the full sample, it is possible that alternative measures could perform better or worse at particular times (though this is made less likely by the fact that coefficients in the alternative models broadly agree). Alternatively, the models could all perform better and worse at different periods. This would in part reflect the specification, where coefficients are intentionally held constant over the full period, leaving the macroeconomic effects on the labour market as a time-varying residual.

Chart 1 compares the four models using the most informative measures of fit: Cramer's λ and RMSE of the aggregate prediction. Both statistics are broken into annual figures by applying the definition of the statistic to each year's data as if it was the full sample. Note that neither the RMSE or Cramer's λ for the full sample is a simple function of the annual figures because there are potentially substantive weighting differences between quarters.

Cramer's λ reflects the prediction performance at the individual level and the pattern evident in the series is one of predictions improving over time. The one exception is the unemployment only model, where performance has levelled off since 1997. This distinction is important as it indicates that more disaggregated models of the available labour force are not only generally better, but this relative performance advantage has improved in recent years. This confirms the importance of allowing for basic information on the composition of the inactive. It leaves the overall ordering of the models intact, while emphasising the disadvantages of looking only at unemployment.

The RMSE of the quarterly aggregate predictions is also shown in Chart 1. The pattern here is quite different from a simple inverse of Cramer's λ. Predictions are most accurate in 1995 for each of these models, although the predictions are also relatively accurate in 1999. Interestingly, the difference between the models is sharpest in the early and later years. In every year, except 1996, the more general models are always favoured over focusing on the unemployment rate alone. The RMSE's of the non-employment index model and the ordinal reasons models are quite similar throughout the period, while the model which uses the detail available on the individuals' reasons for not working has a similar pattern but typically performs slightly worse. Overall, it is generally helpful to predict individual transitions well, but there is a risk of fitting the detailed transitions too well.

Much of the pattern in the RMSE of predictions can be explained by the fact that the model does not pick up all of the cyclical variation. The business cycle was intentionally not accounted for in the models, leaving the individual predictions to be the expected outcome for the mean business cycle conditions. Chart 2 shows the pattern of the average quarterly error so that years of under and over-prediction can be identified. This diagram suggests that there are substantial macroeconomic factors that cause a general over-prediction in the early 1990s, followed by increasing under-prediction that fades after 1997. This is despite the fact the general pattern is towards better predictions in the latest data. It seems unlikely that this pattern is simply a changed desire to take a job on the part of available workers, given that their intentions are informative and increasingly accurate at the individual level.

6.4 Model stability over the sample period

Business cycle effects are general to the labour market and thus should affect all groups of potential workers, although possibly not to the same extent. Other factors could alter the relative transition rates: for example, a rise in student aid or the increased state provision of childcare. The period under consideration also included a number of labour market reforms, notably various 'New Deal' plans to move people out of long-term unemployment and the introduction of the 'National Minimum Wage'.

The estimation strategy used here could in principle be extended to account for these reforms, if the groups that would be affected were easily identifiable. This approach would certainly improve the fit of all of the models, but would diminish the clarity of the separation of variation explained purely by the characteristics of the non-employed. For that reason, we do not use any time-varying controls in the model for the purposes of this paper.

However, it is interesting to see the degree to which variation in the relative probabilities of transition might affect these models. One approach is to estimate the models allowing all transition rates to change annually. This blurs the distinction between aggregate and relative influences, but it should reveal any substantial shifts in the relative transition rates. Chart 3 shows the results of this extension to the ordered reasons model.

The ordered reasons model is interesting because it should reveal any changes in transitions of particularly disadvantaged labour market participants versus those who normally transition at higher rates. These shifts might have occurred due to the labour market reforms, but would also result from a broadening of company schemes to attract non-traditional workers. Shifts in these transition rates would be less evident in the unemployment rate only model because the changes would be diluted in the low transition rate excluded category.

On the basis of the ordered reasons model, it seems that there has been some increase in transition rates but that this is concentrated among the highest transition rate group and occurred mostly before 1997. The limited range of these differences and the unexpected timing illustrates the difficulty of building in 'known policy changes' to this type of measure.[13] While likelihood ratio tests of coefficient stability are rejected with the large sample provided by the Longitudinal LFS, the differences do not appear sufficiently large to alter the aggregate conclusions drawn from these estimates. Overall, while it is technically possible to allow for shifts in relative probabilities within the sample, this extension does not add much to our understanding of the cyclical pattern of labour availability.

7. Evaluating these labour availability estimates

To this point, the analysis has largely focused on producing a statistical model of labour availability, without many specifics about how these estimates could be applied. What should be quite clear from this analysis is that models accounting for the different groups of the non-employed are statistically superior to the standard model that considers only the unemployed and the non-employed. This conclusion is not linked to a particular model of unemployment and so can be treated as a general conclusion.

7.1 The role of differential search rates

There is substantial evidence of differential search rates in the results already shown in Tables 2 and 3. This evidence supports a broader analysis of the matching function, but including other categories of the non-employed might be of limited value if their behaviour over

[13] The fact that the estimates are not significantly different after 1997 should not be interpreted as a test of the effectiveness of recent labour market reforms. For example, the most important New Deal programme over the sample period - introduced in April 1998 - was the New Deal for Young People, which was targeted at 18 to 24 year olds who have been unemployed for over six months. Our analysis considers changes in *average* transition rates, combining the impact on both targeted and untargeted groups, and so cannot tell us about the specific effects of the reforms on their intended target group. For an explicit analysis of the macroeconomic impact of the New Deal for

the business cycle largely paralleled the unemployed but at a consistently lower level of search intensity. In this case, accounting for the additional categories would not alter the time pattern for the overall transitions and thus the unemployment rate would accurately summarise the time series relationship between labour availability and employment growth.

One way to evaluate the importance of differential search intensities in the sample period is to compare the predicted transition rates. Each of the models laid out in Tables 2 and 3 is of individual transition probabilities, so the model's expected transition rates in any period are

$$\text{Transition rate}\{\text{m, t}\} = \sum_{i \in \{\text{non - employed at t}\}} wgt_i F(\mathbf{x}_\mathbf{t}^\mathbf{m}\boldsymbol{\beta}^\mathbf{m} + \mathbf{s}_\mathbf{t}^\mathbf{m}\boldsymbol{\gamma}^\mathbf{m} + \alpha^m) \Big/ \sum_{i \in \{\text{non - employed at t}\}} wgt_i$$

where each model has its own set of $\mathbf{x}_\mathbf{t}^\mathbf{m}$ and $\mathbf{s}_\mathbf{t}^\mathbf{m}$ variables observed at time t and coefficients, but evaluates exactly the same observations. Regardless of each model's simplicity or complexity, every person is assigned an expected transition rate using this approach. Differences in the model predictions are only evident across time because all of the models correctly estimate the average transition rate over the full sample. The number of non-employed people in the denominator is the same for all of the models at time t. These results can be directly compared to evaluate whether the additional information supported by the micro-level data is relevant to the time series pattern.

The average transition rate for the full sample is 0.099. The flat black line in Chart 4 shows this rate. The implied transition rates for four of the alternative models are shown in Chart 4. The importance of differential search rates for understanding the state of the labour market in a particular year is evident from the fact that the expected pattern for the

---

Young People, see Riley and Young (2001).

unemployment model is substantially different from all of the alternatives. The expected transitions from this model account only for the fraction unemployed in the non-employed, which implicitly holds the transition rates for the unemployed and inactive constant. The prediction is primarily due to the variation in the proportion unemployed. The 'unemployment only' model predicts a far sharper decline in the number of workers making themselves available, and the decline continues through 1999. Qualitative differences in the other models are less noticeable, although the full model tends to predict more of a decline in transitions than either the Bank of England non-employment index or the ordered reasons model (which are qualitatively similar in their time patterns).

Each of these models also include the demographic controls, so differences in the coefficients on these factors can also alter the time pattern. Without going through each of the models, Chart 5 shows the typical role demographic factors play in the models. Demographic factors included on their own (with no accounting even for the unemployed/inactive split) do not significantly alter the expected rate of transitions from non-employment from the sample average. Including demographic factors in the ordered reasons model does have a substantial impact, by reducing the slope of the decline. This would occur when demographic shifts occur within the reason categories. The final line in the chart shows (when compared with Chart 4) that the simple unemployment model is also altered by the inclusion of demographics. The line (labelled unemployment only) comes close to looking simply at the unemployment rate as the indicator of labour availability. The sharp decline in this line relative to the other models, combined with earlier goodness-of-fit results, indicate the unemployment rate appears to have exaggerated the decline in labour availability over the sample.

7.2 Using regional variation to evaluate the indices

Without more time series variation, it is hard to develop more business cycle implications of these measures, notably any connection to wage growth. As an exploratory effort, we follow the approach of Bell, Nickell and Quintini (2000) and exploit regional variation in the UK to help identify macroeconomic patterns. In particular, we first explore whether the regional analogues of these measures predict the size of the gross employment inflow differences across regions, which tests whether the availability indices summarise important sources of regional variation in employment growth. This is followed by regressions replacing the unemployment rate with these measures in a regional wage equation.

The regional analogue of the measures is simply the sum of the predicted probabilities within a given region, for example:

$$LA_{rt} = \sum_{i \in \{\text{non - employed at t, in region r}\}} wgt_i F(\mathbf{x_i}\boldsymbol{\beta} + \mathbf{s_i}\boldsymbol{\gamma}' + \alpha)$$

The r subscript refers to a particular region. Realised transition rates (e_{rt}) are a similar sum over non-employed individuals in the region, but the transition variable is 0/1.

The panel regression includes both year and region fixed effects, so that neither the aggregate time pattern nor the constant relative differences between regions identify the response. Specifically, these two equations are estimated:

$$e_{rt} = \beta_1 a_{rt} + \boldsymbol{\delta}_r \mathbf{D}_r + \boldsymbol{\delta}_t \mathbf{D}_t + \upsilon_{rt}$$

$$w_{rt} = \gamma w_{rt-1} + \beta_2 a_{rt} + \boldsymbol{\delta}_r \mathbf{D}_r + \boldsymbol{\delta}_t \mathbf{D}_t + \omega_{rt}$$

where a_{jt} is the regional availability measure under consideration, $\mathbf{D_j}$ and $\mathbf{D_t}$ are vectors of dummy variables for regions and time. The wage equation is estimated with a lagged value, but the results are quite similar if it is estimated for first differences of these wage rates. The regional wage rate is the average wage after including controls for the composition of the workforce as in Bell, Nickell and Quintini (2000).[14] The standard errors for each of these regressions have been corrected for the correlations associated with the panel features.

The coefficient estimates (β_1 and β_2) for these regressions and their p-values are shown in Table 6. For comparison purposes the same equations are estimated with regional claimant count unemployment rate. Claimant count unemployment rates are typically used in regional models because they are less erratic particularly in smaller regions, so we will apply them. The unemployment rate is in different units, so the coefficients are not comparable, but the p-values will reveal the relative effectiveness of the variables. The alternative model availability indices are in comparable units, so their coefficients can be directly compared.

The regressions on the realised transitions rates in the regions reveal that the labour availability measures do help to explain time patterns of taking up jobs in the regions. This might appear to be near a tautology, but at no point is any information that might be characterised as demand for labour within a region included in the model. In addition, regions are not included in the construction of the models, although this could be justified on the basis that labour market statuses are associated with different expectations of getting a job due to differences in benefits take-up rates or other regional variation. Finally, aggregate time variation is removed from the regression by the inclusion of year dummies, so it is not the relative steepness of the measures as revealed in Chart 4 that determines the coefficient values. Interestingly, the model that serves as an analogue to the Bank of England non-employment index does the best job of explaining the regional time variation, although all of the models

[14] The approach to adjusting the regional wages in the earlier figures and throughout this paper follows that of Bell, Nickell and Quintini (2000). Wage equations are estimated on individual data of wages and individual characteristics, with a set of region cross-year effects. These regions cross-year effects estimates become our adjusted mean wage figures for the region.

outperform the unemployment rate. Simply re-expressing the unemployment rate in terms of the expected number of individuals helps to improve the fit, probably because it builds in a constant flow rate from inactivity that will help to account for time patterns in aggregate levels of inactivity.

This is reassuring, but the focus of labour availability is often on wages rates. In these regressions, where the availability rates are directly included, the models do not perform as well as regional unemployment rates. The coefficients on the models are only marginally significant (90% level) at best. This is disappointing, but aggregate fluctuations (including any changes in labour demand) are not accounted for in these regressions. The next section contemplates how to construct appropriate measures of labour market tightness that may be more useful.

7.3 An implicit measure of labour market tightness

Can the availability model be used to generate an indicator of labour market tightness? The job-matching process described and estimated in this paper is consistent with bilateral search, notably a model extended for differential search intensities (Pissarides (2000)). Bilateral search stresses the importance of both sides of the market in evaluating market conditions.

One important result in Pissarides (2000) is that under a range of assumptions these models generate individual decisions that are separable into personal factors and aggregate conditions. The logit models estimated in this paper do not estimate the effects of the aggregate conditions variables identified in the bilateral search model, because these effects vary only across time and the model focused on individual factors. Excluding time-varying parameters results in the estimates *should* result in a systematic pattern of residuals if aggregate factors matter and are time-varying over this sample period. Furthermore, the quarterly average of these residuals is a consistent two-stage estimate of time-specific effects.(15) Estimating the model with time dummies yields the same results but also reveals that the time dummies are jointly significant in these specifications at any conventional levels of certainty.

Treating this time pattern as an estimate of the role of aggregate factors, it makes sense to consider the time effects at a quarterly frequency. The circles in Chart 6 show the quarterly results for the order reasons model. For simplicity, we narrow the focus to this one measure, because the time pattern is quite similar for the non-employment index based model. The pattern shown here parallels the annual figures (Chart 2), but the higher underlying variation in the estimates is revealed. Compared to other time series models this estimation procedure is unusual in that it makes no use of the fact that general labour market conditions in a given quarter are likely to be related to the previous quarter's position. This quarter-to-quarter variation suggests that smoothed estimates of some form probably provide a better measure of underlying tendencies in the labour market. The smoothed line in Chart 6 is the result of LOWESS smoother (with a window of 0.4) being applied to the point estimates.(16)

The problem with evaluating any labour market tightness measure is that there is no universally accepted benchmark. This paper focuses on the weaknesses of the unemployment rate, but each alternative has its own problems. Any direct measure of search-theoretic labour

(15) It is easier to explain the role of these time-varying effects if we assume that they were directly estimated. If dummy variables for the quarters d_t were included, the average marginal effect of being in a particular quarter alters the average marginal effect of a characteristic k according to

$$\left(\overline{M}_k \mid t = t^*\right) = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}^k f(\mathbf{x}_i\hat{\boldsymbol{\beta}} + \mathbf{d_t}\hat{\boldsymbol{\delta}} \mid t = t^*) = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}^k f(\mathbf{x}_i\hat{\boldsymbol{\beta}} + \hat{\delta}_{t=t^*})$$

As long as the effect of individual factors are constants the estimated models maintain the multiplicatively separability of the bilateral search model. This is evident in that the effect of the marginal change in any factor is the coefficient on that factor and the baseline rate, which shifts in response to being in quarter t*.

(16) LOWESS smoother were proposed by Cleveland (1979) to fit trends in data where outliers might be a problem. LOWESS estimates the slope of the smooth curve in the vicinity of each observation (x_i) by employing a weighed regression on all observations within the desired range or bandwidth.

market tightness (*v/u)* relies on highly uncertain measures of the number of actual vacancies ready to be filled. Indeed, the typical measure for the UK—based on vacancies posted with job centres—shows a steady increase throughout the period that is largely indistinguishable from the mirror image of the falling unemployment rate. And survey-based measures of firms' descriptions of the labour market depend both on the nature of question asked and on the details of the sample design (always reflecting only a particular facet of the labour market).

The British Chambers of Commerce surveys (manufacturing and service industries) report the percentage of firms that say they are experiencing difficulties recruiting staff.(17) This is conceptually similar to the aggregate conditions index because it should depend both on the firms' desired levels recruitment and on the availability of labour. On the other hand, it may also be impacted by shortages of particular skills where the relative demand has risen more quickly than the relative supply. The two surveys' results are shown in Chart 7 along with the inverted aggregate conditions index. The pattern in either sector is remarkably similar to the model measure. Another interesting feature of these survey measures is that their longer history can be used to establish a limited measure of what is tight or loose. Over the past twelve years, the manufacturing survey has averaged 55.9, while the services have averaged just below 49.7. Both surveys crossed over their average values in 1995, which suggests this year as the point when labour markets moved to being 'tight', at least relative to the past twelve years.(18)

Applying the same regional wage regression approach used above with these measures of labour market tightness results in a more convincing relationship. Table 8 reports the coefficient and p-value for three labour availability indices and the unemployment rate. Now the alternative

(17) The specific question asked by the BCC is 'Did you experience any difficulties finding suitable staff?' Approximately 2,800 manufacturing and 4,750 service-sector firms answered the survey.
(18) The identification restrictions continue to be relevant. In particular, matching efficiency is assumed constant. In addition, this statement presumes that the tightness implied by the survey responses is stable.

measures all clearly outperform the unemployment rate. Again the Bank of England non-employment index appears to perform particularly well, although the coefficient differences for the alternative labour availability indices are not statistically significant. These differences are entirely driven by the sets of variables included in the alternative models, because the realised transition rates are simply the gross flow to employment from non-employment within each region (which is identical in each alternative model).

This measure of tightness is appealing, but it would require a longer span of data to refine these standard errors and evaluate each model's merits as an indicator. Nonetheless, this work does favour an expanded pool of potential workers when thinking about labour market tightness.

8. Conclusion

Who is willing to work? In the United Kingdom, the answer is a set of people far larger than the ILO unemployed. Evaluating all working-age individuals based on their likelihood of finding a job in three months' time, we find that several categories of the inactive have tendencies to work that equal those of the unemployed. Other categories of the inactive are still relevant for employment, despite their lower transition rates, but they are not direct substitutes for the unemployed. This result had also been shown in the work of Gregg and Wadsworth (1998), but this paper goes beyond that and shows that these groups are meaningful at a business cycle frequency.

How willing are they to work in aggregate? Considering the full set of potential suppliers of labour, this research points toward weighting the non-employed according to their realised transition rates. These transition rates depend on both the reasons/activities of the non-employed and their other characteristics. Both of these factors matter, when predicting who is going to become employed in the next three months and the aggregate number of people becoming employed by the next quarter. This result moves away from the existing literature, which has focused on the creation of simple aggregates to replace the unemployment rate.

Labour market tightness is not just the inverse of availability in a bilateral search model: it also depends on the quantity of workers desired by firms. So the ratio of the number of vacancies to the number of unemployed is a typical measure of labour market tightness. Treating the time-varying component of these availability models as an indication of aggregate conditions yields an alternative 'tightness' measure. The preliminary results shown here indicate that these measures perform reasonably well in regional wage equations.

Overall, the primary limitation of the approach used in this paper is that data availability means that it can only be applied to the United Kingdom over the past eight years. This period includes only expansionary phases of the business cycle. This lack of cyclical history limits both the tests that can be applied to generalised indices of non-employment and the issues to which this type of index can be applied. In the meantime, this approach can help to inform policy-makers seeking to establish the amount of labour services available to the market, at least while the economy continues to expand.

Table 1: Quarterly employment transition rates, by detailed ILO status, 1993-1999

Status	Transition rate to employment	Fraction of non-working population	Fraction of new employment provided
Unemployed	**23.3%**	**23.0%**	**53.5%**
<6 months	34.1	10.5	(35.8)
>=6 months	14.2	12.5	(17.6)
Seeking, unavailable	**23.8**	**2.1**	**5.1**
Student	27.5	0.9	2.5
Looking after family	16.4	0.5	0.8
Sick (temporary)	11.8	0.2	0.2
Sick (long-term)	3.9	0.1	0.0
Other reason	35.2	0.4	1.4
No reason	25.9	0.1	0.2
Not seeking, would like work	**6.6**	**21.3**	**14.1**
Waiting on application	25.2	0.1	0.3
Student	16.4	2.4	3.8
Looking after family	5.6	7.9	4.4
Sick (temporary)	7.6	1.1	0.9
Sick (long-term)	1.3	6.3	0.8
Discouraged	4.3	1.0	0.4
Not started looking	16.4	0.9	1.4
Other reason	12.4	1.7	2.1
No reason	19.0	0.0	0.0
Not seeking, not like work	**5.1**	**53.7**	**27.5**
Waiting on application	29.5	0.0	0.1
Student	13.4	10.8	14.5
Looking after family	3.5	19.3	6.8
Sick (temporary)	4.8	0.8	0.4
Sick (long-term)	0.7	14.1	0.9
Job not needed	3.4	1.8	0.6
Retired	1.9	4.9	0.9
Other reason	8.8	1.6	1.4
No reason	45.6	0.4	1.9

Table 2: Alternative categorisations of availability

Categorisation	Highest availability	→	→	Lowest availability
ILO unemployment	Unemployed, all durations			Inactive, all reasons
Jones & Riddell	Unemployed, all durations	Inactive, all wanting a job		Inactive, no job wanted
Employment Policy Institute	Unemployed, all durations	Inactive, Discouraged + people seeking work, not able to start currently	Inactive, all others wanting a job	Inactive, no job wanted
US Bureau of Labor Statistics alternative measures	Unemployed, 15 weeks or less	Unemployed, remainder	Inactive, Discouraged	Inactive, no job wanted
Bank of England, non-employment index (uses 6 categories)	Unemployed, 6 months or less Unemployed, 6-12 months	Unemployed, remainder	Inactive, Discouraged Inactive, others wanting job	Inactive, no job wanted
Ordinal Reasons	Unemployed, 6 months or less student, seeking other reasons, seeking no reason, seeking waiting for result, like waiting for result, not like no reason, not like	Unemployed, 7 or more months family, seeking short-term sick, seeking student, like not started looking, like not looked, like no reason, like student, not like	family, like short-term sick, like discouraged, like short-term sick, not like other reasons, not like	long-term sick, seeking long-term sick, like long-term sick, not like family, not like no need, not like retired, not like

Table 3: Estimated transition models for detailed status, 1993-1999

Status	Raw transition rates	Marginal, seasonal controls	Marginal, individual characteristics	Marginal, Ordinal Reason var.
Short term unemp.	0.341	0.461	0.312	0.378 ** (0.028)
Student, seeking	0.275	0.430 ♥♥	0.230 ♥	
Other reason, seeking	0.352	0.485 ♠♠	0.331 ♠♠	
No reason, seeking	0.259	0.405 ♥♥	0.258 ♥♥	
Waiting, like	0.252	0.419 ♠	0.274 ♠	
Waiting, not like	0.295	0.458 ♠	0.307 ♠	
No Reason, not like	0.456	0.565 ♠♠	0.411 ♠♠	
Long term unemp.	0.142	0.242	0.149	0.145 ** (0.016)
Family, seeking	0.164	0.316 ♥	0.184 ♦♦	
Short-term sick, seeking	0.118	0.247 ♦	0.141 ♦	
Student, like	0.164	0.305 ♥	0.132	
Not started, like	0.164	0.301 ♥	0.187 ♦♦	
Other reason, like	0.124	0.251 ♦	0.153 ♦♦	
No reason, like	0.190	0.293 ♥	0.183 ♥	
Student, not like	0.134	0.243 ♦	0.095	
Family, like	0.056	0.118	0.041	0.086 ** (0.025)
Short-term sick, like	0.076	0.167	0.087	
Discouraged, like	0.043	0.082	0.051	
Short-term sick, not like	0.048	0.099	0.036	
Other reason, not like	0.088	0.186	0.100	
Long-term sick, like	0.013	-0.030	-0.053	Excluded
Long-term sick, seeking	0.039	0.077	0.021	
Family, not like	0.035	0.058	0.000	
Long-term sick, not like	0.007	-0.063	-0.073	
Job not needed, not like	0.034	0.055	0.014	
Retired, not like	0.019	Excluded	Excluded	
Female			0.002 *	0.005 **
Degree			0.036 **	0.035 **
Higher Vocational Train			0.021 **	0.019 **
A Level			0.003 **	-0.005 **
Vocational Training			0.015 **	0.014 **
Apprenticeship			0.002	0.003 *
Lower Education			-0.028 **	-0.027 **
Age			-0.022 **	-0.015 **
Age^2			6.3E-5 **	4.6E-5 **
Age^3			-5.8E-9 **	-4.5E-9 **
	N= 438,088	lnL=-119,607 N= 438,088	lnL=-117,672 N= 438,088	lnL=-119,820 N= 438,088

♠ indicates coefficient estimate is insignificantly different from the short-term unemployment estimate, ♠♠ indicates significantly greater than the short-term unemployment estimate, ♥ indicates coefficient estimate is insignificantly different from the unemployment estimate, ♥♥ indicates significantly greater than the unemployment estimate, ♦ indicates coefficient estimate is insignificantly different from the long-term unemployment estimate, ♦♦ indicates significantly greater than the long-term unemployment estimate. * and ** indicate coefficient estimate significantly greater than zero, 95 and 99 confidence levels. The seasonal pattern is estimated using 3 quarter dummy variables.

Table 4: Alternative summary regressions on transition to employment

Variable	Ordinal reasons	Unemployment	Unemployment, short and long	Jones & Riddell	Employment Policy Institute	Bank of England Non-employment
Unemployed		0.154 **				
Unemployed, short			0.235 **	0.246 **	0.264 **	0.292 **
Unemployed, med.						0.228 **
Unemployed, long			0.094 **	0.101 **	0.115 **	0.115 **
Order 1	0.378 **					
Order 2	0.145 **					
Order 3	0.086 **					
EPI U2					0.133 **	
EPI U3				0.016 **	0.026 **	
Want a job						0.027 **
Inactive, unavailable						0.157 **
Discouraged						0.032 **
Female	0.005 **	0.003 **	-0.003 **	-0.003 **	-0.001	-0.002
Degree	0.035 **	0.064 **	0.057 **	0.058 **	0.056 **	0.055 **
Higher Vocational Training	0.019 **	0.035 **	0.031 **	0.030 **	0.029 **	0.029 **
A Level	-0.005 **	0.020 **	0.018 **	0.018 **	0.019 **	0.018 **
Vocational Training	0.014 **	0.024 **	0.021 **	0.021 **	0.019 **	0.019 **
Apprenticeship	0.003 *	0.000	0.001	0.000	0.000	-0.000
Lower Education	-0.027 **	-0.037 **	-0.034 **	-0.034 **	-0.034 **	-0.034 **
Age	-0.015 **	-0.041 **	-0.037 **	-0.037 **	-0.037 **	-0.036 **
Age^2	4.6E-04 **	1.1E-04 **	9.9E-05 **	9.9E-05 **	9.8E-5 **	9.5E-5 **
Age^3	-4.5E-09 **	-9.5E-09 **	-8.8E-09 **	-8.7E-09 **	-8.7E-9 **	-8.5E-9**
	LnL=-119,820 n= 438,088	lnL=-124,696 n= 438,088	LnL=-123,194 n= 438,088	lnL=-125,213 n= 438,088	lnL=-122,134 n= 438,088	lnL=-121,890 n= 438,088

Statistical significant coefficients at 90, 95 and 99 confidence levels are noted by one, two and three asterisks respectively. All specifications also include 3 quarter dummy variables to pick up the seasonal pattern in transitions.

Table 5: Likelihood ratio test and goodness of fit for employment transitions

Table 3 models	Likelihood Ratio	Percent Correctly Classified	Cramer's λ	Ben-Akiva & Lemer R^2	RMSE of aggregate flow est.
Base model, no indiv. Vars.	Reject	90.01	0.1224	0.8438	0.0056
Base model	Reject	90.04		0.8461	0.0049
Base model, w/ family variables	Reject	89.63	0.1492	0.8432	0.0055
Ordinal Reason	Reject	90.01	0.1299	0.8453	0.0045
Table 4 models					
Ordinal Reason	Reject	90.01	0.1299	0.8453	0.0045
Unemployment	Reject	89.99	0.0993	0.8398	0.0059
Unemployment, short & long	Reject	89.85	0.1144	0.8423	0.0045
Jones & Riddell	Reject	89.83	0.1145	0.8423	0.0044
Employment Policy Institute	Reject	89.84	0.1192	0.8432	0.0046
Bank of England non-employment	Reject	89.84	0.1216	0.8436	0.0043
Other models					
US Bureau of Labor Statistics	Reject	89.99	0.0635	0.8328	0.0044
Just indiv. vars.	Reject	90.01	0.0522	0.8306	0.0034
Ordinal Reason, no indiv. vars.	Reject	90.01	0.1198	0.8433	0.0054
Unemployment, no indiv. vars.	Reject	90.01	0.0593	0.8310	0.0051
EPI, no indiv.	Reject	90.01	0.0900	0.8365	0.0053
BoE, no indiv.	Reject	90.01	0.0953	0.8376	0.0049

Table 6: Regional implications of labour availability indices

Labour Availability Model	Regional Employment Flow	Regional Wage Rate
Unemployment Rate	-0.0067 (p=0.27)	-0.0249 (p=0.04)
Unemployment model	1.214 (p=0.00)	-0.632 (p=0.07)
Ordinal Reason	1.543 (p=0.00)	-0.598 (p=0.10)
Bank of England non-employment	1.696 (p=0.00)	-0.543 (p=0.14)

Table 7: Regional implications of labour tightness measures

Labour Tightness Model	Regional Wage Rate
Unemployment Rate	-0.0249 (p=0.04)
Unemployment model	11.42 (p=0.01)
Ordinal Reason	15.88 (p=0.00)
Bank of England non-employment	16.24 (p=0.00)

Chart 1: Goodness of fit measures for alternative models

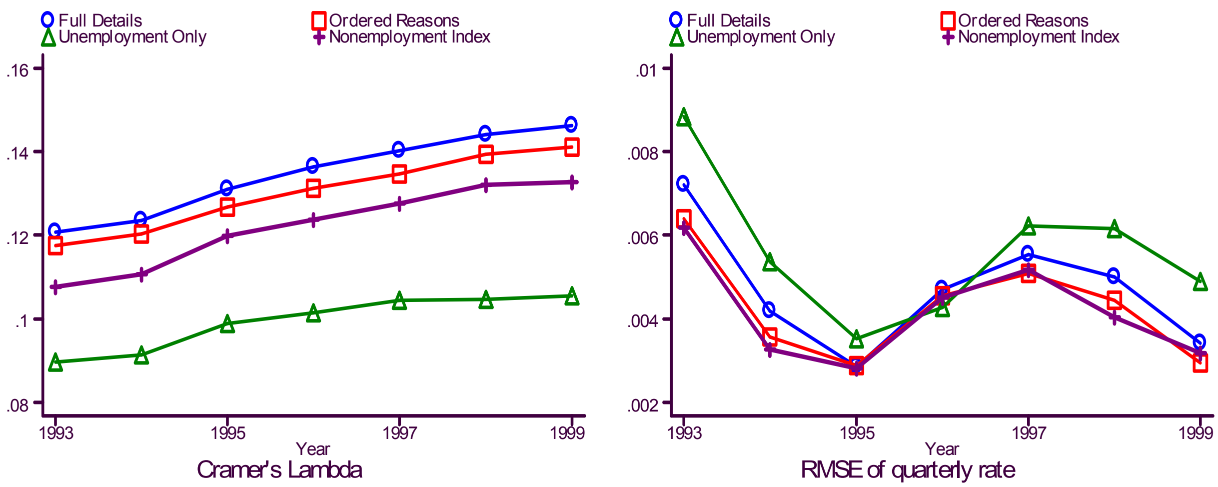


Chart 2: Average annual errors (predictions-actual)

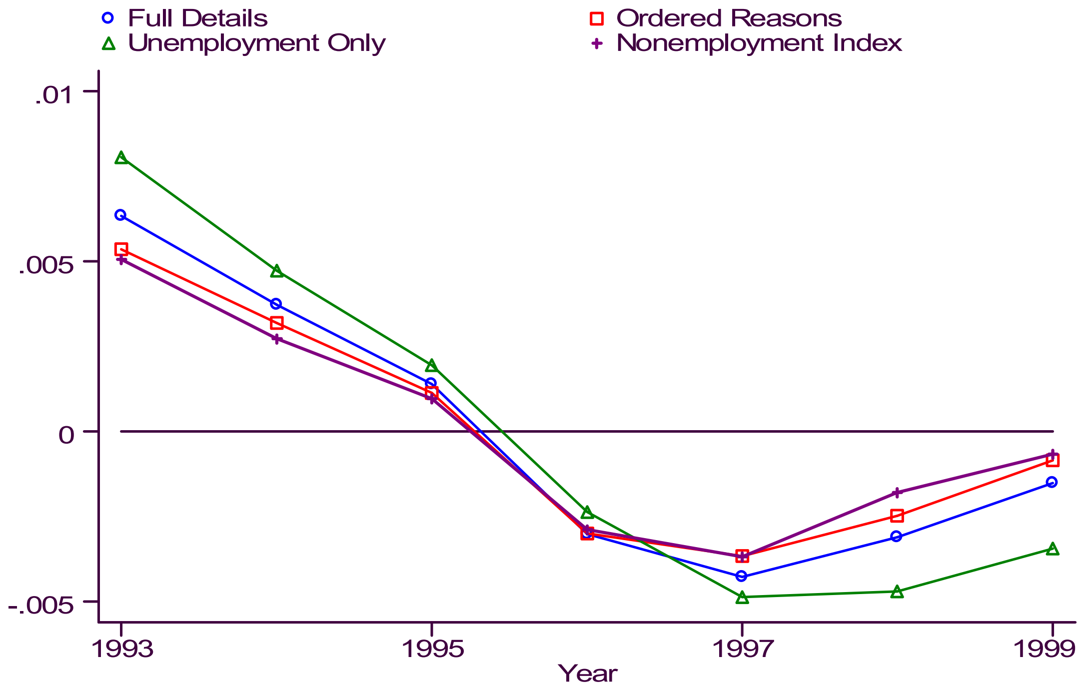


Chart 3: Predicted transition rates in less-restricted ordered reasons model

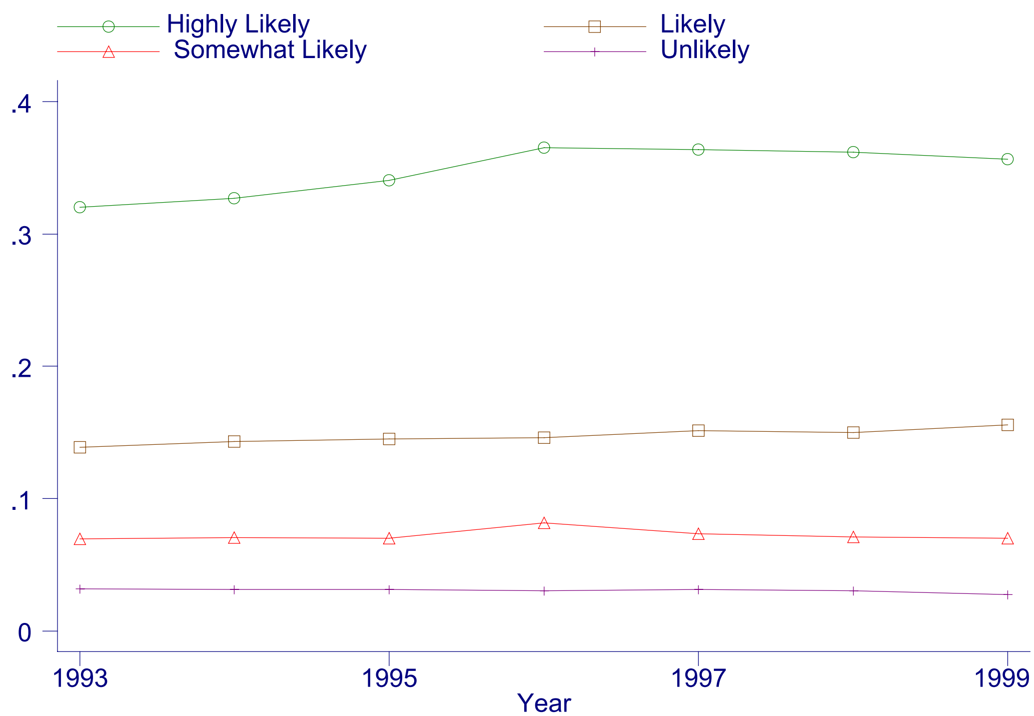


Chart 4: Model comparisons, predicted transition rates

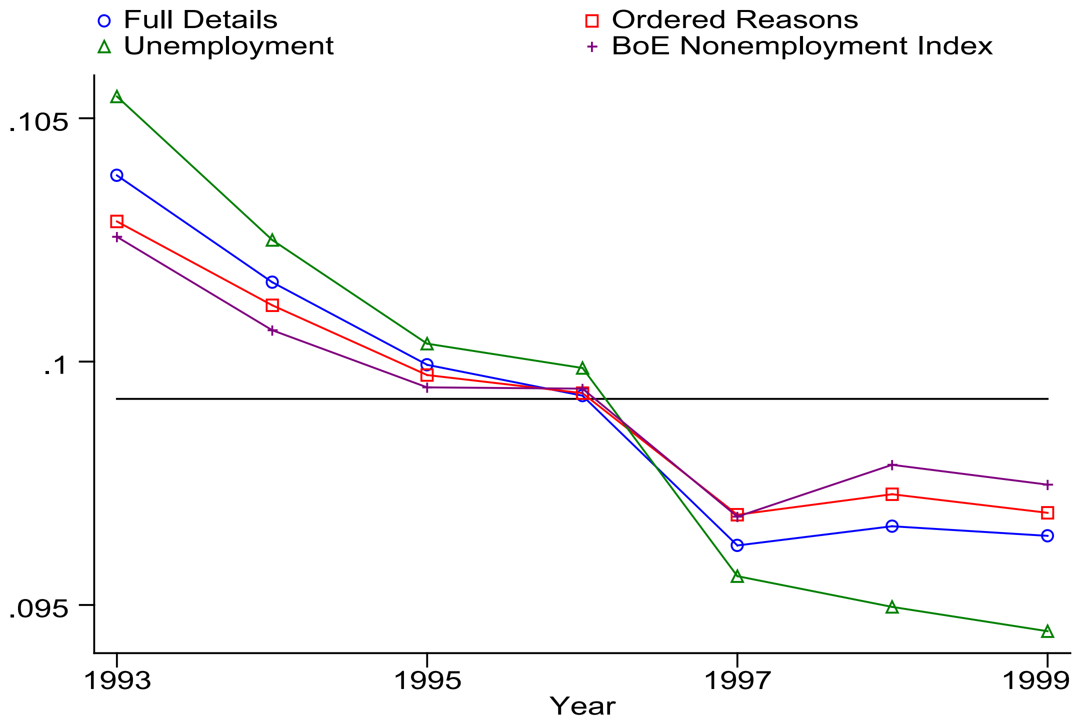


Chart 5: Contributions to ordered reason model of transition rate

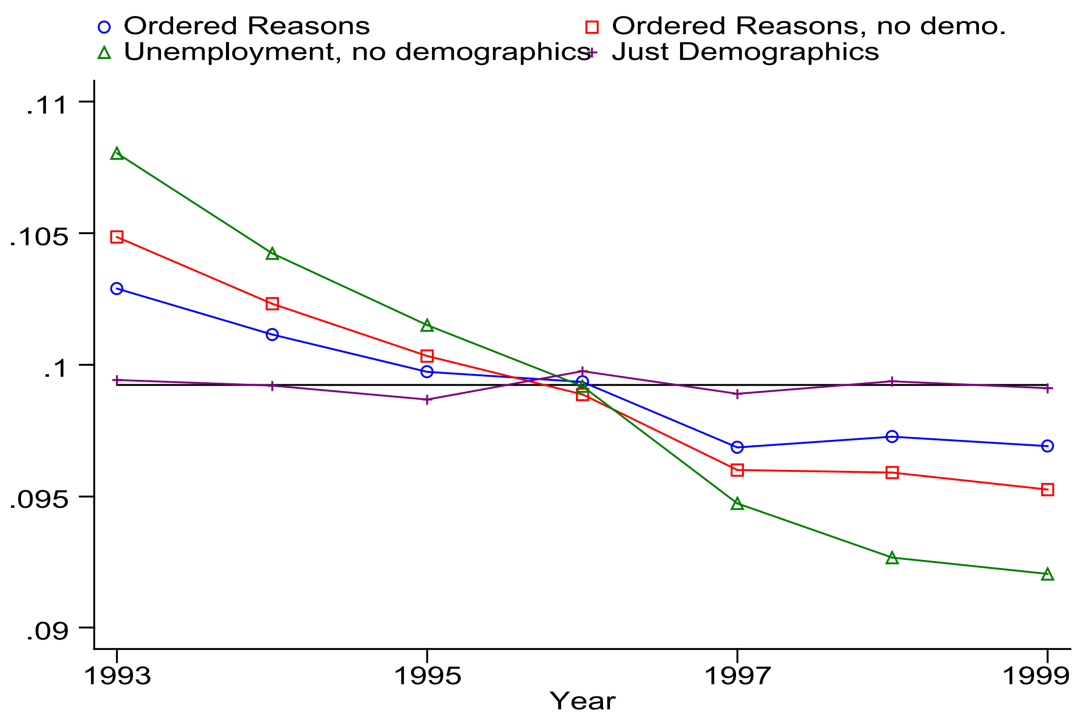


Chart 6: Implied index of aggregate conditions

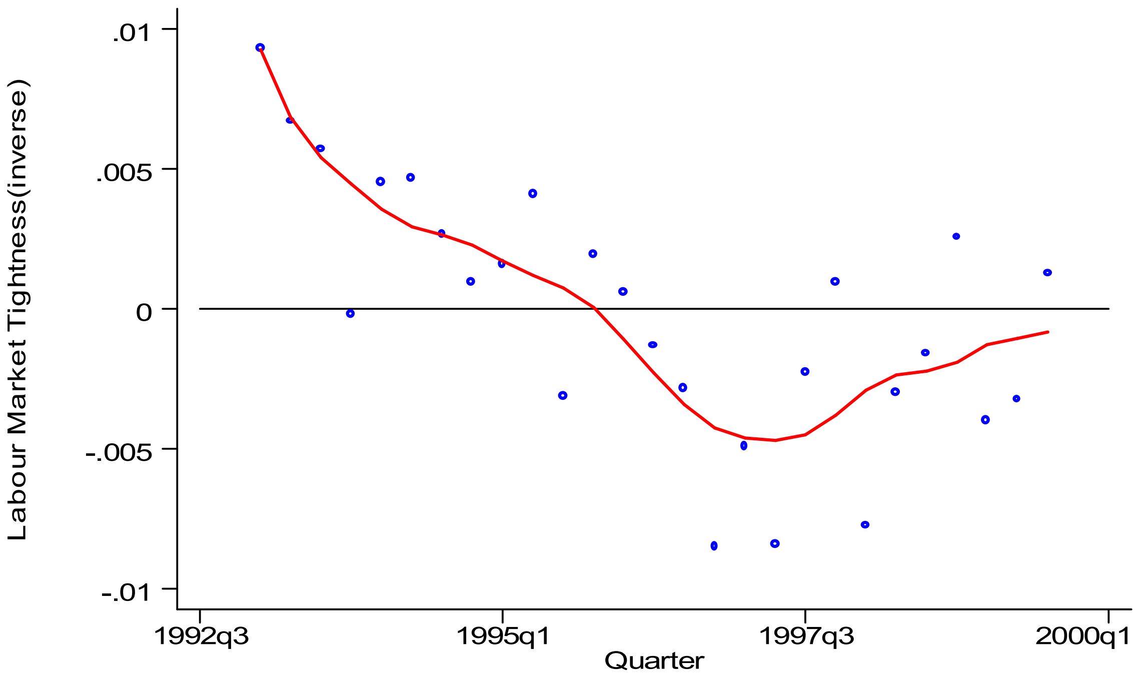
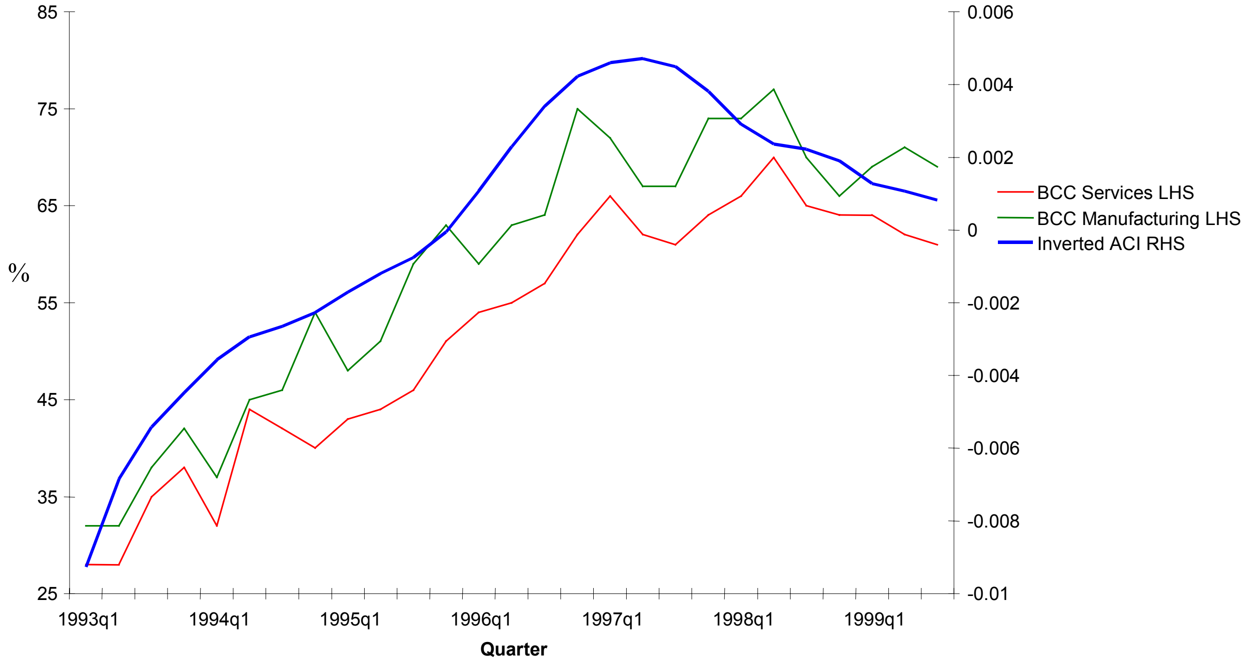


Chart 7: Aggregate conditions versus survey measures of recruitment difficulties



85
75
65
55
45
35
25
%
0.006
0.004
0.002
0
-0.002
-0.004
-0.006
-0.008
-0.01
1993q1
1994q1
1995q1
1996q1
1997q1
1998q1
1999q1
Quarter
BCC Services LHS
BCC Manufacturing LHS
Inverted ACI RHS

References

Bell, B, Nickell, S and Quintini, G (2000), 'Wage equations, wage curves, all that', *Centre for Economic Performance Discussion Paper*, DP0427.

Ben-Akiva, M and Lerman, S (1985), *Discrete choice analysis*, London: MIT Press.

Brigden, A and Thomas, J (2000), 'What does theory tell us about labour market tightness?', manuscript.

Clark, K and Summers, L H (1979), 'Labour market dynamics and unemployment: a reconsideration', *Brookings Papers on Economic Activity*, Vol. 1979:1, pages 13-60.

Cleveland, W S (1979), 'Robust locally weighted regression and smoothing scatter plots', *Journal of the American Statistical Association*, Vol. 79, pages 829-36.

Cramer, J (1999), 'Predictive performance of the binary logit model in unbalanced samples', *Journal of the Royal Statistical Society, Series D (The Statistician)*, Vol. 44, pages 85-94.

Employment Policy Institute (1999), *Employment audit*, Issue 10.

Flinn, C and Heckman, J (1983), 'Are unemployment and out of the labor force behaviourally distinct states', *Journal of Labor Economics*, Vol. 1, No. 1, pages 28-42.

Gilleskie, D B and Mroz, T A (2000), 'Estimating the effects of covariates on health expenditures', manuscript.

Gregg, P and Wadsworth, J (1998), 'Unemployment and non-employment: unpacking economic inactivity', Employment Policy Institute *Economic Report*.

Jones, S R G and Riddell, W C (1999), 'The measurement of unemployment: an empirical approach', *Econometrica*, Vol. 67, No. 1, pages 147-62.

King, M (1998), 'Monetary policy and the labour market', *Economic Report* (Employment Policy Institute), Vol. 12, No. 9, pages 1-15.

National Statistics (2000), 'Release of labour force survey longitudinal datasets', *Labour Market Trends*, page 528.

Nickell, S (2001), 'Has the UK labour market performance changed', speech at the Society of Business Economists, 16 May 2001.

Pissarides, C A (2000), *Equilibrium unemployment theory*, Cambridge, Massachusetts: MIT Press.

Riley, R and Young, G (2001), 'The macroeconomic impact of the New Deal for Young People', *NIESR Discussion Paper no. 184*.

Shiskin, J (1976), 'Employment and unemployment: the doughnut or the hole?', *Monthly Labor Review*, pages 3-10.

Tano, D (1991), 'Are unemployment and out of the labor force behaviourally distinct labor market states? New evidence from the gross change data', *Economic Letters*, Vol. 36, pages 113-17.

Appendix

Per cent correctly classified

Conventionally, this is evaluated at the point where the prediction probability for the individual ($\hat{F}_i$) is above 0.5:

$$\text{Per cent correctly classified} = \frac{1}{n}\sum_{i=1}^{n}\Big(y_i \mid (\hat{F}_i \geq 0.5)\Big) + \Big((1 - y_i) \mid (\hat{F}_i < 0.5)\Big)$$

where $y_i = 1$ when an individual transition is made and zero otherwise.

Ben-Akiva and Lerman R-squared

Ben-Akiva and Lerman (1985) suggest a measure based on the average probability of correct prediction.

$$R^2_{BL} = \frac{1}{n}\sum_{i=1}^{n} y_i \hat{F}_i + (1 - y_i)(1 - \hat{F}_i)$$

Cramer's Lambda

In Cramer's Lambda the typical forecast error for each outcome is given equal weight regardless of the relative frequencies of the outcomes.

$$\lambda = (\text{average } \hat{F}_i \,|\, y_i = 1) - (\text{average } \hat{F}_i \,|\, y_i = 0)$$

Root Mean Squared Error

The RMSE of this model is based on the aggregate quarterly predictions of the model.

Expressed in the notation used for the individual level fit measures this is:

$$RMSE = \sqrt{\sum_{t=1}^{T}\left(\sum_{i\in t}\frac{\hat{F}_i}{n_t} - \sum_{i\in t}\frac{y_i}{n_t}\right)^2}$$

Federal Reserve Bank
of Cleveland
Research Department
P.O. Box 6387
Cleveland, OH 44101